IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPEAL FROM THE EXAMINER TO THE BOARD OF PATENT APPEALS AND INTERFERENCES

Application No.:		09/707,624
Applicant:	:	Overton et al.
Filed	:	November 7, 2000
Title:	:	Method and Apparatus for Real Time Insertion of Images Into Video
Art Unit	:	2614
Examiner	:	Desir, Jean Wicel
Docket	:	131235.1001
Customer No	:	32914

MAIL STOP: APPEAL BRIEF PATENTS
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Dear Sir:

APPEAL BRIEF

Applicants have appealed to the Board of Patent Appeals and Interferences from the decision of the Examiner mailed November 2, 2005, finally rejecting Claims 17-29. Applicants filed a Request for Pre-Appeal Brief Review together with a Notice of Appeal on April 3, 2006.

TABLE OF CONTENTS

REAL PARTY IN INTEREST

The present application was assigned to Mirage Systems, Inc., by an assignment from the inventor recorded on November 7, 2000, in the Assignment Records of the United States Patent and Trademark Office at Reel 011275, Frame 0887. The present application was subsequently assigned to Vistas Unlimited, Inc., by an assignment from Mirage Systems, Inc., recorded on July 16, 2002, in the Assignment Records of the United States Patent and Trademark Office at Reel 013084, Frame 0301.

RELATED APPEALS AND INTERFERENCES

There are no known appeals or interferences that will directly affect or be directly affected by or have a bearing on the Board's decision in this pending appeal.

STATUS OF CLAIMS

Claims 1-29 are pending in the application; claims 1-16 are allowed, and claims 17-29 are rejected pursuant to a Final Office Action mailed November 2, 2005. Claims 17-29 are presented for appeal.

STATUS OF AMENDMENTS

No amendment has been filed subsequent to the mailing of the Final Office Action.

SUMMARY OF CLAIMED SUBJECT MATTER

The subject matter of claims 17-29 generally pertain to methods and apparatus for inserting images into video taken of a real site by a camera. An example of a use of such methods and apparatus is in inserting in real-time advertising into video of a sporting event.

In the method of independent claim 17, each image in an image stream from the camera is processed as follows. A background area within the original image is identified from at least one target area in which a target image will be inserted. A target image is rendered based on at least a predefined three-dimensional model of the at least one target area within the site and the camera's position and pointing direction. The three-dimensional model does not model the entire site. Occlusions are then identified within the at least one target area of the original image. The background of the original image, the at least one target image, and the image of the occlusions are then combined into an output image.

In the apparatus of independent claim 28, a model renderer generates a "synthetic image" based on a predefined three-dimensional model of a target area within a site from a known position of a camera. The three-dimensional model is of less than the entire site. The synthetic image includes a target image inserted in the target area. An image separator separates the background from the target area and creates a masked background image. An image combiner combines the masked background image with the target image.

The three dimension model refers to a computer model, such as a CAD model, which describes certain elements of the venue comprising the model to a predefined three-dimensional coordinate system. See, e.g., page 3, lines 17-20 of the specification. These elements include the at least one target area, where images will be inserted. This model is then rendered from the perspective of the camera to generate a target image.

GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

1. Whether the rejection of claims 17-29 under 35 U.S.C. § 102(b), as being anticipated by U.S. Patent No. 5,491,517 of Kreitman et al. ("Kreitman et al.") was in error.

ARGUMENT

A. Standard

1. 35 U.S.C. § 102(e)

Under 35 U.S.C. § 102(e), a claim is anticipated only if each and every element as set forth in the claim is found in a single prior art reference. *Verdegaal Bros. v. Union Oil Co. of California*, 814 F.2d 628, 631 (Fed. Cir. 1987); M.P.E.P. § 2131. In addition, "[t]he identical invention must be shown in as complete detail as contained in the . . . claims" and "[t]he elements must be arranged as required by the claim." *Richardson v. Suzuki Motor Co.*, 9. U.S.P.Q.2d 1913, 1920 (Fed. Cir. 1989); *In re Bond*, 15 U.S.P.Q.2d 1566 (Fed. Cir. 1990); M.P.E.P. § 2131.

B. Argument

1. Rejection of Claims 17-29 Under § 102(b) is in error

It is respectfully submitted that the rejection of claims 17-29, as being anticipated by US Patent No. 5,491,517 of Kreitman et al., is in error.

a. Independent claims 17 and 28

Claims 17 and 28 require a "pre-defined three-dimensional model" of a target area of the site, and rendering or generating a target image based on that model. The examiner cites for both claims elements 62 and 64 of Figure 5, Figure 2, Figure 21, column 13, lines 26-51, and column 14, lines 46-55 are cited as disclosing this limitation. In paragraph 4 of the final Office action dated November 2, 2005, the examiner admits that Kreitman et al. do not mention a predefined three-dimensional model, but argues that such a model is inherent. Applicants respectfully agree that it is not mentioned, but disagree that it is inherent.

Figure 2 is simply an illustration of the venue, showing placement of the cameras. See col. 3, lines 54-56 ("FIG. 2 is a schematic illustration of a tennis game used as an example for explaining the operation of the system of FIG. 1.")

Figure 5 references a "model" available to a perspective identification unit 62. Element 62 is described at lines 30-36 of column 7 as "identifying the viewing angle and zoom of an active camera 30 and for determining an appropriate perspective

transformation between the model 50 and the input video frame." Model 50 is shown in Figure 3, which the examiner does not reference, is a "geometric model" of the tennis court 32, "typically as a top view." Figure 3 plainly illustrates only a 2 dimensional model.

The long passage cited in column 13 describes use of a perspective matrix for a camera, which includes information on the position, rotation angles and the zoom of the camera. There is no mention of a three dimensional model being used. Figure 21 illustrates a table for storing camera parameters, namely the position, current title, angles and zoom, and a matching coeffecient. The cited passage in column 14 describes determining a perspective transformation matrix based on the current values of the camera and a line and point identified in a frame and associated with the model.

As already mentioned, the examiner has admitted that Kreitman does not mention 3-dimensional model "as claimed," but argues that such a model is "inherent to the Kreitman's disclosure." His reasoning was that:

> ... the target image or synthetic image is based on a predefined three-dimensional (3-D) reference model as claimed, because the camera (30 of Fig. 2) is pointed to a site which is a tennis game; the tennis game is based on a 3-D reference model of a real field (site) where the game is taking place... .
>
> Final Office action, November 2, 2005, paragraph 4.

The examiner's reasoning in support of his inherence argument is simply not coherent. He admits that there is no 3-D model disclosed by Kreitman et al., but then says that "tennis game is based on a 3-D reference model of the real field." The tennis game is certainly being played in three dimensions, but the statement that the tennis game based on a 3-D model is nonsensical. It is respectfully submitted that the examiner has failed to set out reasoning that supports his contention of inherency.

More importantly, Kreitman et al. cannot inherently disclose use of a 3-D model of the site. Their system and methods can only work with a 2-dimensional model. Figure 3 shows this model with the image that is to be inserted. It is plainly 2 dimensional and corresponds to the top view of the tennis court. To use this model a perspective transformation is made between the model and the video frame. They state: "The system 14 also comprises a transformer 64 for transforming the implantation data from the model plane to the image viewing plan ... ." Column 7,

lines 36-38 (emphasis supplied). Implantation data refers to the image that will be inserted, an image location mask and a blending mask. Figure 15 illustrates the concept of the transformation matrix. It clearly contemplates only planar objects. A "feature extraction unit extracts features of interest from the images – for example, lines on the tennis court – and a "perspective identification unit" then compares the extracted features with the model and produces a transformation matrix. Col. 7, lines 57-63. There is no mention of the possibility of using 3 dimensional models that applicants can find, and there is no discussion of how such models could be used.

The subject matter being discussed in columns 13 and 14, which the examiner relies on, concerns a method of calibrating cameras so that camera parameters – pan, tilt and zoom – can be used to create a transformation matrix. This calibration takes place after each new "cut." Col. 13, lines 53-55. The method requires identification of lines and points in the images and comparison of those lines and points to the planar surface model of the court. Thus, the model must include the lines of the court. Perspective transformation matrices for the cameras bring a line identified in a frame of the video into alignment with the 2 dimensional model of the tennis court. Column 14, lines 45-48. The perspective matrix is calculated based on the camera position values. The perspective matrix is then used to determine a transformation matrix M for transformation of the model (element 50 of Figure 3) into the plane of the image in the frame. Column 14, lines 56-61. The lines in the model of the court and the lines in the image frame are then compared to create a "match value." *Id.* This process is repeated for a different camera, with a different position coordinate, if the match value is poor for other lines and points identified in the image frame. Col. 15, lines 1-4. The transformation matrix with the best match is then used in transformation matrix to render the image to be inserted in the correct perspective.

The claimed invention, on the other hand, utilizes a fundamentally different process. It begins with a three dimensional model of the site of the target area, in which the image is to be inserted, has been defined. For each frame the target image is rendered from the model in 2-dimensions from the perspective of the camera, based on parametric data from the camera for the particular frame. There are a number of advantages to this process as compared to process in Kreitman et al., including the

ability to define the target image in three dimensions, as well as on surfaces that do not exist at the venue.

In short, Kreitman et al. does not and cannot meet, either literally or inherently, each and every limitation of the two independent claims for at least these reasons. Therefore, they cannot anticipate the independent claims or any of the claims depending from them.

b. Dependent claim 18

Claim 18 requires identifying background area within the original frame by identifying at least one target area based on a predefined model of selected surfaces of the site and the camera's position and pointing direction. The examiner cites element 60 of Figure 5 as meeting this limitation. However, the model is plainly shown in the figure being used only by the perspective identification unit 62, which is subsequent to the element 60 (the feature identification unit). There is no suggestion of using the model to generate a background mask that applicants can find. Therefore, the rejection of this claim is in error for at least this reason.

c. Dependent claim 19

Claim 19 requires, among other things, rendering a model and then generating from the rendering a mask defining the target area. The figure and passage cited by the examiner, Figure 21 and column 13, lines 26-51, concern only a perspective matrix. As previously explained, this is used to create a transformation matrix for transforming the plane of the 2-D model into the plane of the image. This is not rendering a three dimensional model. The rejection of this claim is therefore in error for at least this reason.

d. Dependent claims 21 and 23

Claims 21 and 23 require that the camera's position be defined within the model. However, the 2-D model of Kreitman et al. does not include the positions of the camera. Figure 3, which illustrates the model, does not show them. Col. 13, lines 37-53 and col. 3, lines 33-40 are cited as disclosing this limitation in claim 23;

column 13, lines 26-55 and Figure 21 are cited as disclosing this same limitation in connection with claim 21.

Column 13, lines 26-55 concern using camera position information to shorten the process of creating a perspective matrix. No mention is made of the camera being part of the model. Indeed, since the camera is being used to determine the matrix that is used to transform the plane of the model into the plane of the image, applicants do not see how it could be, or why it would be, part of the model.

For at least these reasons, Kreitman et al. cannot anticipate claims 21 and 23, and therefore the rejections of claims 21 and 23 are in error.

e. Dependent claim 29

Claim 29 requires, among other things, "an occlusion separator for comparing the reference image of the second synthetic image to the target area of the video image and generating an image of occlusions within the target area of the video image...." An occlusion is an element in the foreground, such as a player, that partially obstructs the surface (real or imaginary) on which an image is inserted. The image cannot be inserted over the player. The claim requires separating the occlusions by comparing a reference image to the target area of the actual video.

The examiner cites elements 60 (a "feature identification unit") and 62 (a "perspective identification unit") of Figure 5, as well as Figure 3 and 4A-4C. According to the description of Figure 5, the perspective identification unit 62 identifies the viewing angle and zoom of the camera based on features in video frames identified by feature identification unit 60, and determines an appropriate perspective transformation between the model 50 and the input video frame. Column 7, lines 30-36. Feature identification unit 60 produces a background mask that is used by the mixer for creating a background into which the transformed image is implanted, which allows the combined output to take into account occlusions. Column 8, lines 1-6. However, according to Figure 5, this mask is generated from feature identification unit 60, which does not make use of the model. Figures 4A-4C illustrate example masks prepared by a designer. Column 6, lines 17-18, lines 20-21 and lines 28-29.

There appears to applicants to be no suggestion of rendering a synthetic reference image based on a predefined 3-dimensional model, and comparing it to original frame to separate out occlusions within the target area. Kreitman et al. cannot anticipate claim 29 for at least this reason, and therefore the rejection is in error.

2. Other Errors

Given the fundamental errors addressed above, applicants have chosen not to address other errors in the examiner's reasoning. The failure to address them should not be construed as acquiescence in the examiner's reasoning.

CONCLUSION

In view of the errors noted above in the examiner's rejections of claims 17-29, applicants respectfully request the Board of Patent Appeals and Interferences to reverse the final rejection of the examiner and instruct the examiner to issue a notice of allowance of all claims.

The Commissioner is hereby authorized to charge any fees or credit any overpayments to Deposit Account No. 07-0153 of Gardere Wynne Sewell LLP, referencing docket number 131235.1001.

Respectfully submitted,

/Marc A. Hubbard/

Marc A. Hubbard
Registration No. 32,506

Date: November 3, 2006

Correspondence To:

Gardere Wynne Sewell LLP
1601 Elm St, Suite 3000
Dallas, Texas 75201
(214) 999-3000

CLAIMS APPENDIX

17. A method for inserting a target image into an image stream of a real site taken by a camera, the method comprising:

receiving a stream of original images taken by a camera of a real site;

for each original image in the image stream,

identifying a background area within the original image from at least one target area in which a target image will be inserted;

rendering, for each at least one target area, a target image based on at least a predefined three-dimensional model of the at least one target area within the site and the camera's position and pointing direction, the three-dimensional model being of less than the entire site;

identifying occlusions within the at least one target area of the original image; and

combining the background of the original image, the at least one target image, and the image of the occlusions into an output image.

18. The method of claim 17, wherein identifying a background area within the original image includes identifying at least one target area within the image based on a predefined model of selected surfaces of the site and the camera's position and pointing direction.

19. The method of claim 18, wherein identifying at least one target area within the image comprises:

receiving camera parameters, the camera parameters including parameters indicating the pointing direction of the camera;

rendering the predefined model from the camera's position, based on the camera parameters; and

generating from the rendering a mask or defining the target area within the image.

20. The method of claim 19, wherein the camera parameters included parameters indicating the camera's position within the site.

21. The method of claim 19, wherein the camera's position is predefined within the model.

22. The method of claim 17, wherein rendering, for at least one target area, a target image comprises:

receiving camera parameters, the camera parameters including parameters indicating the pointing direction of the camera; and

rendering the predefined model based at least in part on the camera parameters.

23. The method of claim 22, wherein the camera's position is predefined within the model.

24. The method of claim 22, wherein the camera parameters includes parameters indicating the camera's position within the site.

25. The method of clai m 24 wherein the camera parameters are encoded onto a video signal generating by the camera.

26. The method of claim 17 wherein identifying occlusions with the image of the target area includes:

rendering the predefined model of the site with a reference image inserted in each of the at least one target areas; and

comparing the reference image to the target area within original image.

27. The method of claim 17 further including encoding the output image and camera parameters onto an output video signal, the camera parameters including parameters indicating the pointing direction of the camera.

28. An apparatus for inserting a target image into video of a real site taken by a camera, the apparatus comprising:

a model renderer for generating, a synthetic image based on a predefined three-dimensional reference model including a target area within a site from a known

position of a camera, the three-dimensional model being of less than the entire site, the synthetic image having a target I mage inserted in the target area;

an image separator for masking from a video image contained within a frame of a video signal generated by the camera, the target area to create a masked background image; and

an image combiner for combining the masked background image with the target image.

29. The apparatus of claim 28 further comprising:

a second model renderer for generating a second synthetic image based on the predefined reference model of the site, the second synthetic image having a reference image inserted into the target area; and

an occlusion separator for comparing the reference image of the second synthetic image to the target area of the video image and generating an image of occlusions within the target area of the video image;

wherein the combiner further combines the image of the occlusions with the masked background image and the target image.

EVIDENCE APPENDIX

None

RELATED PROCEEDINGS APPENDIX

None